

CLAIMS

What is claimed is:

1. 1. A method in a computerized device for maintaining a client session in a network having a plurality of routers, the network having an application executed at a plurality of replicas, comprising the steps of:
 - 4 providing a database of bindings of request identifiers to replicas where each binding is a record having a request identifier, a replica identifier and a binding expiration time, the database associated with a first router of the plurality of routers;
 - 7 maintaining a change log of records entered into the database, each change log entry having a change event generated by the first router and an event number sequential to an event number of a preceding change event in the change log;
 - 10 maintaining a current version vector associated with the database and the change log, the current version vector entry for the first router being a most recent event number from the change log, the current version vector entry for each other router being a most recent event number received at the first router from that other router;
 - 14 receiving an update of change events generated at another router in the plurality;
 - 15 reconciling the current version vector according to the received update; and
 - 16 reconciling the database according to the received update such that the client session is maintained.
- 1 2. The method of claim 1 wherein the request identifier is a client identifier and an application identifier.
- 1 3. The method of claim 2 wherein the client identifier is an Internet Protocol address.
- 1 4. The method of claim 2 wherein the client identifier is a dproxy Internet Protocol address such that the binding associates a dproxy with a replica.

1 5. The method of claim 1 wherein the step of reconciling the current version vector
2 comprises the steps of:

3 comparing a least recent event number of the router that generated the update to
4 the event number in the current version vector entry for that router;

5 if the least recent event number is in series with the event numbers in the database
6 as determined by the current version vector entry for that other router, then entering the
7 most recent event number of the received update into the current version vector entry for
8 the router that generated the update of change events; and

9 if the least recent event number in the update is not in succession to the event
10 number in the current version vector entry for the router that generated the update of
11 change events, then discarding the received update.

12

1 6. The method of claim 5 wherein the step of reconciling the database further comprises:

2 if the update was not discarded in the step of reconciling the current version
3 vector, then for each entry of the received update,

4 a) determining whether the received entry has expired;

5 b) if the received entry has expired, then discarding the entry;

6 c) if the received entry has not expired, then comparing the request identifier of
7 the received entry with the request identifier in the entries in the database;

8 d) if a matching entry is not found for the received entry, adding the received
9 entry to the database;

10 e) if a matching entry is found for the received entry, then comparing the
11 application identifier of the received entry with the application identifier of the matching
12 entry;

13 f) if the application identifiers match, then retaining the entry having a later
14 expiration time in the database; and

15 g) if the application identifiers do not match, then retaining an entry selected
16 based on a deterministic function applied to a portion of each entry.

1 7. The method of claim 6 wherein the step of retaining an entry based on a deterministic
2 function comprises the steps of applying a function to the application identifiers; and
3 selecting an entry based on the outcome of the function.

1 8. The method of claim 6 wherein the step of retaining an entry based on a deterministic
2 function comprises the steps of applying the deterministic function to the request
3 identifier; and selecting an entry based on the outcome of the deterministic function.

1 9. The method of claim 1 further comprising the step of deleting a binding from the
2 database when the expiration time for the binding has been exceeded .

1 10. The method of claim 1 further comprising the step of sending a request for an update
2 of change events to another router in the plurality; and
3 the step of receiving the update further comprises receiving the update in response
4 to the request.

1 11. The method of claim 1 further comprising the steps of:
2 periodically generating a first router update of change events; and,
3 transmitting the first router update of change events to at least one other router in
4 the plurality.

1 12. The method of claim 1 further comprising the steps of:
2 affirming that an update has been received from each router of the plurality within
3 a predetermined period for each router;
4 if an update has not been received from a router within the predetermined period
5 for that router, requesting an update of change events from that router; and
6 if an update is received in response to the request,
7 reconciling the current version vector according to the received update;
8 and
9 reconciling the database according to the received update.

1 13. The method of claim 5 wherein the step of reconciling the database further comprises
2 the steps of:

3 determining from the received update whether the database has a complete record
4 of changes based on the current version vector;

5 if the database does not have a complete record of changes , requesting a
6 replacement database from a router of the plurality of routers.

1 14. The method of claim 1 further comprising the step of transmitting a copy of the
2 database and the current version vector to another router of the plurality of routers in
3 response to a request from the other router.

1 15. The method of claim 1 wherein the computerized device fails temporarily and
2 recovers, the method further comprising the steps of:

3 writing the first router change log to a persistent storage device;

4 sending an update of change events written to the change log in the persistent
5 storage device to other routers in the network;

6 after recovering from failure, requesting a database and an associated version
7 vector from one of the routers in the plurality;

8 retaining the received database and associated version vector;

9 reconciling the received database with the change log from the persistent storage
10 device; and

11 updating the received version vector.

1 16. The method of claim 1 wherein the received update includes a version vector and the
2 method of maintaining a current version vector further comprises the step of maintaining
3 the current version vector in a version vector table including past version vectors.

1 17. The method of claim 16 further comprising the steps of:

2 determining from the version vector table whether the database is current based
3 on the version vector table; and
4 if the database is not current, then requesting missed change events from a second
5 router in the network .

1 18. The method of claim 17 wherein each router caches updates received from other
2 routers in the plurality, the method further comprising the step of:
3 if the router that generated the received update does not respond to the request for
4 missed change events, requesting the missed change events from a second router of the
5 plurality and reconciling the change events into the database and current version vector.

1 19. The method of claim 1 wherein the computerized device fails temporarily and
2 recovers, wherein the step of maintaining a current version vector further comprises the
3 steps of:

4 creating an epoch timestamp from a clock of the computerized device to mark a
5 recovery period;
6 entering a value pair to the current version vector for the first router, the value
7 pair being an event number and the epoch timestamp; and
8 the method further comprising the step of after recovery, requesting a database
9 copy and associated version vector from one of the other routers in the plurality.

1 20. The method of claim 19 further comprising the steps of :

2 determining whether a pre-selected time period has passed; and
3 deleting value pairs before a most recent value pair from the current version
4 vector having timestamps created before the pre-selected time period.

1 21. A system to maintain a client session in a network having a plurality of routers, the
2 network having an application executing at a plurality of replicas, comprising:
3 a network interface to receive a request from a client to access the application;

4 a storage device to store a database of bindings of requests to replicas where each
5 binding is a record of a request identifier, a replica identifier and a binding expiration
6 time, the storage device to further store a change log and a current version vector
7 associated with the change log where the change log includes records added to the
8 database by the system and the current version vector has an entry for each router in the
9 network, each entry storing a version vector of a particular router; and

10 a controller coupled to the interface and the storage device, the controller
11 configured to route the request, if the controller finds a binding matching the request, the
12 controller to route the request to a replica of the plurality according to the binding and to
13 reset the binding expiration time, if the controller does not find a binding matching the
14 request, the controller to add a new record to the database having information from the
15 request to form a binding of the client to a replica of the plurality of replicas.

1 22. The system of claim 21 wherein the network interface is configured to receive an
2 update of change events for the database from another router in the plurality of routers in
3 the network; and the controller is further configured to reconcile the database according
4 to the received update and to update the current version vector in response to reconciling
5 the database.

1 23. The system of claim 21 wherein the controller is configured to transmit periodically,
2 to at least one of the other routers in the plurality, an update of change events and the
3 current version vector.

1 24. A computerized device to maintain a client session in a network having a plurality of
2 such computerized devices, the network having an application executing at a plurality of
3 replicas, comprising:

4 means for providing a database of bindings of request identifiers to replicas where
5 each binding is a record having a request identifier, a replica identifier and a binding
6 expiration time, the database associated with a first router of the plurality of routers;

7 means for maintaining a change log of records entered into the database, each
8 change log entry having a change event generated by the first router and an event number
9 sequential to an event number of a preceding change event in the change log;

10 means for maintaining a current version vector associated with the database and
11 the change log, the current version vector entry for the first router being a most recent
12 event number from the change log, the current version vector entry for each other router
13 being a most recent event number received at the first router from that other router;

14 means for receiving an update of change events generated at another router in the
15 plurality;

16 means for reconciling the current version vector according to the received update;
17 and

18 means for reconciling the database according to the received update such that the
19 client session is maintained.

1 25. A computer program product having a computer-readable medium including
2 computer program logic encoded thereon that, when performed on a computer system
3 having a coupling of a memory, a processor, and at least one communications interface,
4 provides a method for maintaining a client session in a network having an application
5 executing at a plurality of replicas by performing the operations of:

6 providing a database of bindings of request identifiers to replicas where each
7 binding is a record having a request identifier, a replica identifier and a binding
8 expiration time, the database associated with a first router of the plurality of routers;

9 maintaining a change log of records entered into the database, each change log
10 entry having a change event generated by the first router and an event number sequential
11 to an event number of a preceding change event in the change log;

12 maintaining a current version vector associated with the database and the change
13 log, the current version vector entry for the first router being a most recent event number
14 from the change log, the current version vector entry for each other router being a most
15 recent event number received at the first router from that other router;

16 receiving an update of change events generated at another router in the plurality;

17 reconciling the current version vector according to the received update; and
18 reconciling the database according to the received update such that the client
19 session is maintained.

1 26. A method in a computerized device for maintaining a client session in a network
2 having a plurality of routers, the network having an application executed at a plurality of
3 replicas, comprising the steps of:

4 providing a database of bindings of requests from clients to replicas where each
5 binding is a record having a request from a client, a replica identifier and a binding
6 expiration time, the database associated with a first router of the plurality of routers;

7 maintaining a change log of records entered into the database, each change log
8 entry having a change event generated by the first router and an event number sequential
9 to an event number of a preceding change event in the change log;

10 maintaining a current version vector associated with the database and the change
11 log, the current version vector entry for the first router being a most recent event number
12 from the change log, the current version vector entry for each other router being a most
13 recent event number received at the first router from that other router;

14 receiving a request from the client to access the application, the request having a
15 client identifier and an application identifier;

16 comparing data from the request with records in the database;

17 if the database has a record with a client identifier and an application identifier
18 matching the data from the request,

19 a) routing the request to a replica of the plurality of replicas according to the
20 matching record; and

21 b) resetting the binding expiration time of the matching record; and

22 if the database does not have a record with a client identifier and a replica
23 identifier matching the data from the request,

24 a) routing the request to one of the plurality of replicas; and

25 b) entering, into the database a new record forming a binding of the request to one
26 of the replicas.

- 1 27. The method of claim 26 wherein the router is a DNS server and wherein the
- 2 application identifier in the request is a domain name and wherein the step of routing
- 3 comprises mapping the request to an Internet Protocol address of the one replica.